DOCUMENT RESUME

ED 304 769 EA 020 781

AUTHOR Gabriel, Roy M.; Langhorst, Beth Hoover
TITLE Uses of the Regional Database, 1987-88. OERI Database and School Profiling Project.
INSTITUTION Northwest Regional Educational Lab., Portland, Oreg.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE Nov 88
CONTRACT 400-86-0006
NOTE 34p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Database Design; *Database Management Systems; Database Producers; *Databases; Elementary Secondary Education; *Information Retrieval; *Information Storage; Online Vendors; Research Tools; *School Statistics
IDENTIFIERS *Northwest Regional Educational Laboratory

ABSTRACT

This report describes and illustrates the Northwest Regional Laboratory's (NWREL) Database and School Profiling project. Uses and implications for future data assembly and database design activities are explained. Analysis of the database has impacted evaluation and policymaking activities regarding at-risk youth, rural education, early childhood education, American Indian education, and local strategic planning activities. Two groups compose the database's content. The first group contains data from the U.S. Census on 1,300 school districts in the Northwest classified according to district size, rurality, and poverty; ethnic composition of districts; dropout statistics; and educational attainment of adults in a district. The second group contains data from state assessment and evaluation files in the Northwest region according to student achievement, attitudes, and behavior; enrollment and attendance; participation in special programs; and teacher experience. The conclusion lists lessons learned, which include difficulties in data gathering such as outdated data and merging data from different sources. It also mentions possible directions for the future, which include using a commercial database. An appendix defines the indicators used in the Oregon district files. (JAM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

177-310 LA
(225)

ED304769

Uses of the Regional Database 1987–88

Roy M. Gabriel

Beth Hoover Langhorst

OERI Database and School Profiling Project
Northwest Regional Educational Laboratory
Portland, Oregon

November, 1988

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

This paper is based upon work performed pursuant to Contract 400-86-0006 of the Office of Educational Research and Improvement of the United States Department of Education. It does not, however, reflect the views of that agency or the Northwest Regional Educational Laboratory.

EA 020 781

BEST COPY AVAILABLE

Table of Contents

	Page
Acknowledgments	ii
Abstract	iii
Introduction	1
Initial Development of the Regional Database	3
Evolution and Further Uses of the Regional Database: 1987-88	
National Use: The Condition of Small, Rural Schools	6
Regional Use: Federal Aid to Schools with Large Indian Populations	10
State Use: Early Childhood Education	12
State Use: Oregon Business and Education Partnerships	15
Local Use: District Profiles	17
Lessons Learned and Future Directions	21
References	23
Appendices	24

Acknowledgments

The development, enhancements and uses of NWREL's regional database have profitted from the insights and efforts of several staff at the Laboratory in addition to the authors of this summary report. Dr. Patricia Anderson, the previous director of the OERI Database project, is most responsible for the initiation and early development of the database. She identified this demonstration effort as high priority among an otherwise indistinguishable set of potential directions for the project. Phil Nickel provided much of the early and painful quality control and file assembly of a sometimes baffling array of "foreign" data sets. He also did much of the development of the district profile described in this report. Finally, Dennis Deck lent his considerable skill and experience to many of the database design issues we continue to grapple with as the regional database moves beyond a demonstration effort and into a fully implemented one.

Abstract

At the suggestion of its advisory group, NWREL's Database and School Profiling project initiated the development of a regional database in 1986. Acquiring data from federal and state sources, a focus on the design and early implementation issues soon gave way to a variety of requests for information at a range of levels: national, regional, state and local. Analysis of the database have impacted evaluation and policymaking activities regarding at-risk youth, rural education, early childhood education, Indian education, and local strategic planning activities. In this report, the authors attempt to describe and illustrate these uses and their implications for the project's future data assembly and database design activities.

INTRODUCTION

The Database and School Profiling project at the Northwest Regional Educational Laboratory began in December, 1985 with heavy interest, the burden of great potential, but little concerted direction for project activities. Under a general theme of assisting state and local educators to make better use of data and information in their decision and policy making activities, the project had at least the following activities in store:

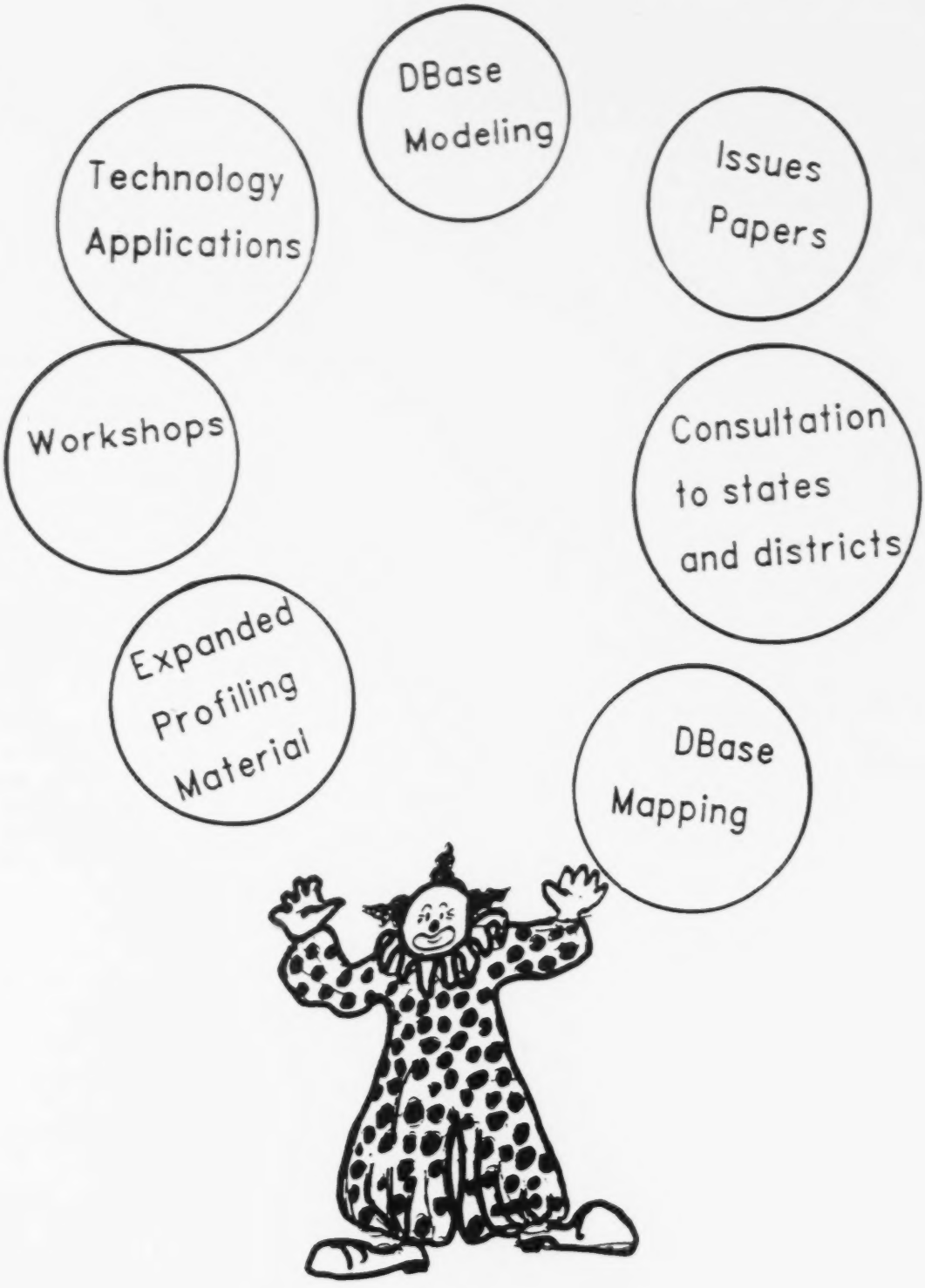
- Expanded Profiling Material
- Workshops
- Technology Applications
- DBase Modeling
- Issue Papers
- Consultation to states and districts
- DBase Mapping

While these all had the interest and support of varying regional constituencies, the limited resources of the project demanded greater focus. Such a focus was slow in coming, and in the initial year of the project, NWREL staff associated with its efforts often felt like the clown pictured in Figure 1, juggling too many balls in the air, but fearing that dropping any one of them would disrupt the rhythm of the act and the audience would lose interest.

Declining resources allocated to the project lent greater importance to and indeed forced the issue of priority setting late in the project's first year. Conferences with Northwest assessment directors and advisory groups to NWREL's Evaluation and Assessment program provided useful direction. Since several of the intended activities related to database design and implementation issues, these clients urged us to meet these objectives with an in-depth, long term demonstration effort. They suggested that project staff develop a region-wide database that would provide both useful information to educators in all Northwest states and a detailed illustration of the process involved in the development and use of a large scale database of information.

In this report the authors attempt to describe the process of this initial development and the variety of subsequent enhancements related to requests for information of the database. Issues addressed in the initial construction of the database are followed by illustrations of the use of the informatiion at federal, regional, state and local levels. Concluding remarks include lessons learned, and future directions for the "second generation" regional database.

Figure 1. Juggling Competing Interests in the Initial Project Year



INITIAL DEVELOPMENT OF THE REGIONAL DATABASE

In responding to the direction suggested by the project's advisory groups, at least two major issues needed resolution. The first related to the content or issue focus of the database. Project staff felt strongly that the development of the database required some specification of topic or educational issues to drive appropriate data gathering activities. The response to this was immediate and universal among project advisors. The needs of at-risk youth were paramount in the minds of educators around the Northwest.

The second primary issue concerned the level at which data were needed. That is, were regional or state-level summaries sufficient? Or were data at progressively more specific levels required: school district, school building or individual student? This decision had major implications for the source of the information gathered and for the level of effort required of the project. Again, the advisory group and project staff reached consensus quickly. The school district was the essential unit of specificity. State-level summaries were too general. They would not allow any understanding of the variations within a state. There were some 1,300 districts in the six states in the Northwest region. Data at the building or student level was not recommended for reasons of availability and magnitude of effort.

Information needs were gleaned by project staff. Demographic characteristics, financial resources and student achievement indicators were of most interest. These necessitated two levels of data collection. The initial interest in tapping solely federal sources (e.g., U.S. Census, Center for Education Statistics, Market Data Retrieval) would not be sufficient. While district demographics and financial resources could be obtained, information on student achievement could not. The Northwest states, who had expressed interest in this demonstration, had to supply the district-level achievement data.

Initial assembly of the database progressed over a six month period. U.S. census tapes were obtained from the Center for Education Statistics. Data extracted from these files comprised a "common core" of information for the six Northwest states, i.e., all states had the same indicators and common definitions of terms. This was an important advantage of this data source. When providing a regional depiction of poverty or enrollment, for example, one did not have to qualify the findings by describing the different measurement approaches, times at which data were collected, etc. Advantages of uniformity and comprehensiveness were balanced by the recency of the data, however. Census data collected in 1980 could have questionable validity to advise decision-making in 1987.

State-specific data were obtained from the six states in the Northwest. In addition to student achievement, NWREL obtained a variety of other demographic and financial indicators to cross-validate the census data collected seven years earlier. Tapping what were essentially six different sources of information posed difficulties in the time needed to acquire the data and the effort needed to process it. Data from states came in many forms and variations. It was a major staff effort to read, interpret, verify and adequately document the data sent by the six states. Yet, it provided information available from no other, more central, source, and the data submitted was far more recent than that available from the U.S. census.

A sample of the indicators extracted from these federal and state sources and merged on the regional database is presented in Table 1. For the first time, NWREL could provide its clients with data-based information on questions ranging from simple to complex:

- How many school districts are there in the region?
- What proportion of them have enrollments over 25,000?
- What proportion of these districts have more than 20% of its students from families in poverty?
- What is the distribution of minority enrollment in any given state?
- What is the dropout rate in each state?
- Does the dropout rate differ for districts of varying size?

Table 1. Database and School Profiling Project:
Contents of Regional Data Base

Data from U.S. Census on 1300 school districts in the Northwest

- District size, Rurality and Poverty
- Ethnic Composition of Districts
- Dropout Statistics
- Educational Attainment of Adults in District

Data from State Assessment and Evaluation File in Northwest Region

- Student Achievement, Attitudes and Behavior
- Enrollment and Attendance
- Participation in Special Programs
- Teacher Experience

These and other questions like them were addressed in a report designed to describe the regional database and illustrate its potential applications (Gabriel & Anderson, 1987). One of the tables in that report is reprinted here to further demonstrate the initial uses of the database. In Table 2, the distribution of dropout rates across districts of varying size and rurality are shown for two Northwestern states.

Table 2. Range of Dropout Rates within Districts of Similar Size and Rurality in Idaho and Oregon

Size/Rurality:	10% or Less	11% to 20%	21% to 30%	31%to 40%	41% to 50%	51% or More
State 1						
Very Small Rural	24	21	13	3	0	2
Small Rural	0	11	0	0	0	0
Large Rural	0	2	0	0	0	0
Small Not Rural	0	9	6	0	0	0
Large Not Rural	4	11	4	0	0	0
V Large Not Rural	0	1	0	0	0	0
State 2						
Very Small Rural	25	41	10	2	1	0
Small Rural	0	13	2	0	0	0
Large Rural	0	4	1	0	0	0
Small Not Rural	5	12	10	1	0	0
Large Not Rural	7	33	5	0	0	0
V Large Not Rural	1	1	1	0	0	0

Very Small = Fewer than 1,000 students
Small = Between 1,000 and 2,500 students
Large = Between 2,500 and 20,000 students
Very Large = More than 20,000 students

Rural = More than 75% of the children living in rural settings
Not Rural = Less than 75% of the children living in rural settings

THE EVOLUTION AND FURTHER USES OF THE REGIONAL DATABASE: 1987-88

The initial focus of the regional database--issues related to at-risk youth throughout the region--began to shift in relation to additional requests for its application to other policy issues. These uses are described in this section, organized by their level of application: national, state or local.

NATIONAL USES OF THE DATABASE

National Rural, Small Schools Task Force Report to the Regional Educational Laboratories

Need for data gathering and analysis

The purpose of the data gathering and analysis described in the following paragraphs was to present to the National Rural, Small Schools Task Force a picture of how rural, small schools were represented on demographic, financial and performance indicators.

Data gathering activities

For the purpose of this project it was necessary to expand the existing regional at-risk youth database. Demographic, financial and performance indicators were identified that would address most directly the issues pertaining to rural, small schools. These were then contrasted with what was available in the NWREL database and what additional information was reasonable to acquire from individual states. Appropriate demographic indicators were selected from the Census portion of the regional database. Financial indicators were identified from the information available from Center for Educational Statistics Common Core Data tapes. Additional financial data and student achievement indicators were needed from the states themselves.

A national sample of seventeen states was selected and individuals in each state were identified as contact persons to retrieve achievement and Chapter 1 funding and participation information. Achievement data were received from nine of the seventeen states.

In an effort to determine how rural, small schools are represented among the most needy schools in each state, the following indicators were selected:

Family Wealth

- median family income
- percent of families with children below poverty level

School Wealth

- per pupil expenditure by district

Student Performance

- achievement data from statewide assessment
- percent "disconnected" youth (age 16-19 who were not in school, not working and not in the labor force)

Other indicators for each district included the following:

- size (enrollment)
- percent rurality of district population
- percent of revenue obtained from local, state, federal sources

Analysis and Sample Results

Demographic and financial data for each state was downloaded from mainframe computer files to create SPSS-PC files on a microcomputer. This made it easier to add data obtained on hardcopy. A simple database was set up for hardcopy data entry and, when entry was completed the data were exported and merged with the SPSS-PC files.

Districts in each state were assigned to one of six mutually exclusive groups based on enrollment (Very Small = < 500, Small = < 2500, and Not Small) and the percent of the population living in rural areas (Rural = at least 75% of the population live in a town with less than 2,500 population or in an unincorporated area). This yielded six demographic categories of districts in each state:

Very Small, Rural	Very Small, Not Rural
Small, Rural	Small, Not Rural
Not Small, Rural	Not Small, Not Rural

Descriptive statistical analyses were conducted for each of the seventeen states separately. Frequency distributions of all financial and achievement indicators were run within each state and cutoff points for the four intrastate quartiles on each indicator were established. The distribution of each of the six groups of districts described above across these quartiles was calculated for each indicator. If rural, small schools exhibited no greater needs than other schools, it would be expected that approximately 25% of these schools would fall in each quartile of the distribution on each indicator. Since quartiles were established separately for each state, interstate differences (e.g., in funding patterns) did not contaminate these comparisons.

For purposes of the National Task Force report, the district was considered to fall into an extreme quartile (bottom or top) if either one of the indicators in a category fell into the bottom or top quartile. The number and percent of districts falling in the bottom intrastate quartile, aggregated across states is shown in Table 3.

Table 3. Number and Percent of Districts by Size/Rurality in Bottom Quartile of Key Indicators

DISTRICT CATEGORY:	INDICATOR: Family Wealth	School Wealth	Student Performance	Total
Very Small, Rural				
Number	772	584	538	2337
Percent	33%	25%	23%	
Small, Rural				
Number	66	67	65	257
Percent	26%	26%	26%	
All Other				
Number	338	533	474	2131
Percent	16%	25%	22%	

Disproportionality is shown most clearly in Family Wealth. About one-third of very small, rural schools are in the bottom quartile of their state in terms of median family income and/or the percent of families with children living below the poverty level. In contrast, school district which are not small and not rural find only sixteen percent in this disadvantaged condition -- about half the proportion of very small, rural schools. In terms of what they put into students' education (expenditures per pupil) and what they get out of it (student performance), the districts are fairly equivalent.

A second analysis sought to identify that proportion of districts that exhibited multiple indicators of "poorness". For each size/rurality category the number and percentage of districts falling in the bottom quartile of the state on one, two, or all three indicators were summarized. For the final report, the Very Small and Small categories were combined. Also reported was the number of children (sum of enrollment across districts) in each category. Table 4. contains the results of this analysis for one of the states in the sample.

Table 4. Outcome for Rural School Districts

Number of Rural Districts	322
Rural Districts with Poor Student Outcomes	10
Rural Districts with Poor Family Income	139
Rural Districts with Poor Per Pupil Expenditure	79
Rural Districts Poor on 2 or More Indicators	27
(Children enrolled in 27 Districts)	14,427
Small, Rural Districts Poor on 3 or More Indicators	1
(Children enrolled in the 1 District)	468

In reviewing data such as these across the entire sample, a notion of "chronic poorness" was developed to describe those districts in the bottom quartile in their states on two or more indicators. The number of rural districts, small and large, and the number of students enrolled in these districts is shown in Table 5. Also shown are the projected totals for the entire 50 states.

Table 5. Rural School Districts with Two or More Indicators of Poorness

	17 States *(Sample)*	50 States *(Projected)*
Small, Rural, Poor School Districts	717	2,280
Large, Rural, Poor School Districts	66	470
Total Rural, Poor School Districts	783	2,750
Children Attending Small, Rural, Poor School Districts	316,386	1,252,734
Children Attending Large, Rural, Poor School Districts	314,549	983,024
Total Children Attending Rural, Poor School Districts	*630,935*	*2,235,758*

Uses of analyses

These data and results were eventually utilized in a report to Congress in support of the Hawkins-Stafford Elementary and Secondary Educational Amendments of 1988, creating a new Rural Educational Opportunities program.

REGIONAL USES OF THE DATABASE

Report to the Indian Education Policy Board, a Northwest Regional Educational Laboratory Advisory Committee

Need for data gathering and analysis

The purpose of the data gathering and analysis described in the following paragraphs was to present to the Indian Education Policy Board a picture of how schools in Alaska, Idaho, Montana, Oregon and Washington with an Indian population of more than 10% were represented on demographic, financial and performance indicators.

Data gathering activities

Data from the existing regional at-risk youth database were used as the basis of information for this project. The same demographic, financial and performance indicators used for the Rural Education analysis presented earlier were output to the new database for the states of Alaska, Idaho, Montana, Oregon and Washington. Data on Indian enrollment in individual districts and schools were obtained from state sources.

Process of analysis

A simple database was set up for each state with data from the regional database imported for each district. Information from state records on the percentage of Indian enrollment was used to identify districts with schools having Indian populations of 10% or more. This new database was created to reside and be maintained by the Indian Education program at NWREL.

Analysis and Sample Results

Analysis conducted in this application were similar to that of the Rural Education application. Within-state quartiles were established on eight indicators for all six states in the region. School districts having any school with more than 10% Indian students enrolled were identified in each state and included in a tabular representation such as that shown in Table 6. Presence in the bottom or top quartile of the state on any indicator signified the district's "distressed" (*) or "healthy" (#) condition, respectively, on that indicator. Table 6. is a sample of this analysis for one Northwest state.

Uses of analysis

The analysis was presented to the Indian Education Policy Board, a NWREL advisory committee.

Table 6. Preliminary Report for Indian Schools: Distressed and Healthy Schools

School District	Med Fam Inc	Fam Bel Pov	Dis con Yth	Pct Loc Fnds	Pct State Fnds	Pct Fed Fnds	Per Pup Exp	Comb Stu Ach	Num Ind Bldg	Num Dis Bldg	Percent Native Enrollment
	#			15	80	5			1	15	3
		#		7	69	24	#	#	1	2	56
	*	*	#	4	75	20	#		1	1	42
	*	*		4	82	14			3	3	12
			*	13	78	9			5	8	12
				3	91	6	#		2	2	23
			#	6	74	20		#	4	4	27
	*	*	*	2	76	23			1	1	28
	*		*	3	66	31	#	#	2	2	83
	#			12	84	4	*		1	2	12
	*	*	#	2	51	46	#	#	1	1	73
				11	67	22			3	3	31
		*	*	4	80	15			3	3	23
	#	#		11	83	6			2	11	6
	#	#	*	15	33	52	#	#	1	1	19
	*		*	3	57	40		#	3	3	70
	#		*	8	87	5	#		1	4	9
	*	*	*	4	41	55	#	#	1	1	20
				12	80	8	*		3	1	9
		*		13	75	12			2	4	19
	#	#	#	8	87	6	*		1	6	6
		*		2	80	18		#	2	2	24
	*			8	84	8			1	4	9
			*	8	82	10	*		2	5	8
			*	7	76	17			6	6	28
				10	80	8	*		1	10	5
	*	*	#	8	63	30	#	#	1	1	59
		*	*	6	87	7		#	1	2	10
	#			21	69	9	#	#	2	50	3
			*	7	82	8	*		1	6	7
		#		12	82	5			3	3	14
		*		16	77	7			4	56	3
		*		19	72	8	#		1	62	2
	*	*	*	6	48	46	#	#	1	1	23
	*	*	*	2	72	25	#		6	6	18
	*	*		1	72	27	#		5	5	30
		*		1	56	42		#	2	2	00
				12	77	11		#	2	2	14
		#		13	81	6			1	5	6
TOTALS DISTRESSED CATEGORIES	12	16	14	--	--	--	6	0			
TOTALS HEALTHY CATEGORIES	7	6	5	--	--	--	13	15			
AVERAGE FUNDING DISTRIBUTION	--	--	--	8	73	18	--	--			

Note 1: The asterisks (*) and pound signs (#) listed below Median Family Income, Families Below Poverty, Disconnected Youth and Per Pupil Expenditure are markers for quartile rankings. The asterisk (*) is an indicator of a distressed school.

STATE USES OF THE DATABASE

Sample Site Selection for the State of Washington, Department of Community Development, Early Childhood Education and Assistance Program Longitudinal Evaluation Study

Need for data gathering and analysis

The Washington Early Childhood Education and Assistance Program (ECEAP) funds sites throughout the state to provide educational and family assistance to 4-year-old, socio-economically at-risk children. The purpose of the data gathering and analysis described in the following paragraphs was to provide additional information on poverty and rurality for each local program site in order to balance these factors in the sample selection process of the longitudinal evaluation study.

Data gathering activities

Data from the existing regional database at NWREL were used as the basis of information for this project. Poverty indicators and percent rurality information were produced for every school district associated with each ECEAP local program site. Other information on the specific programs throughout the state was obtained from ECEAP evaluation activities.

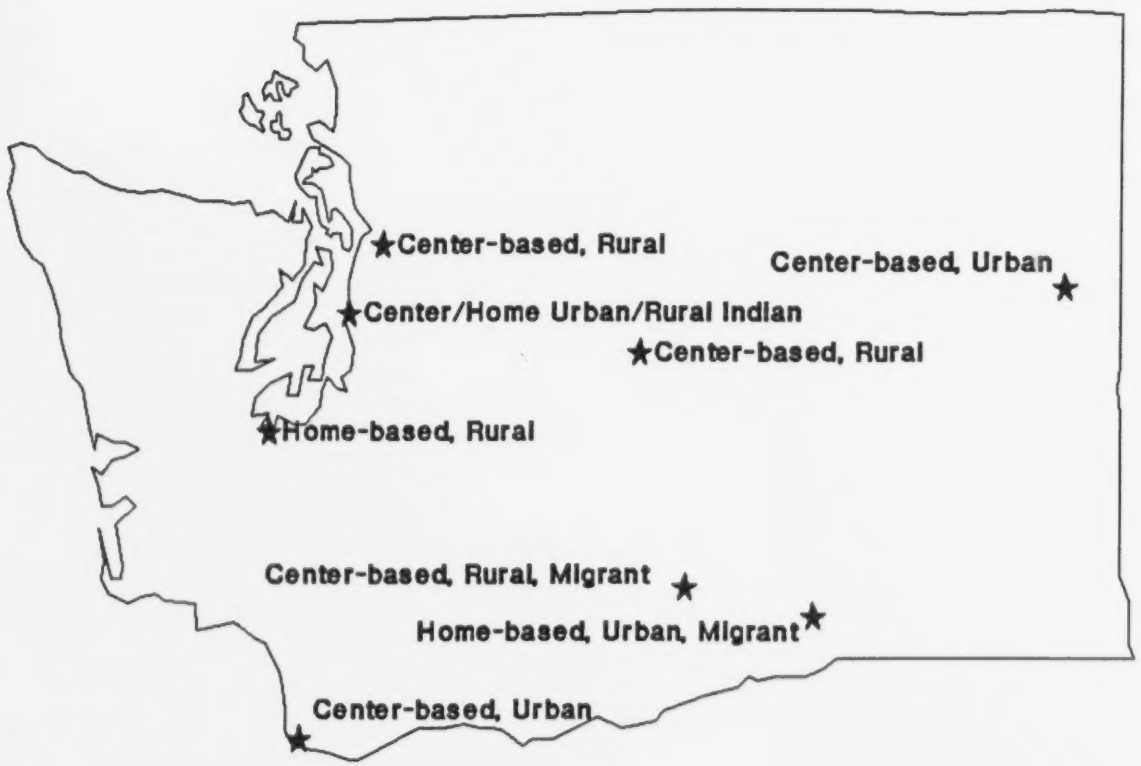
Analysis and Sample Results

Poverty and rurality factors were added to local program characteristics (delineated below) to form a matrix from which one third of the sites were selected for each of the first three initial sampling years. Programs were grouped into "cells" by program model (center- vs. home-based), location (urban, rural), population (non-minority, migrant, Indian) sponsoring organization, and the amount of parent participation required. Program sites were then randomly assigned to a data collection year (or "wave") so that these factors were balanced in each year. Table 7. shows the selection for the first year. Figure 2. illustrates the geographical distribution of the first year sample site selection.

Table 7. Distribution of ECEAP Sample Selection Criteria

	No. Programs	No. Selected
Program Type:		
Center-based	13	4
Combined	4	2
Home-based	4	2
Location:		
Urban	10	3
Combined	7	4
Rural	4	1
Population		
Non-minority	11	5
& Migrant	4	2
& Indian	2	1
& Both	4	
Organization		
School District	6	3
Non-Profit Org.	7	2
Comm. Action Prog.	4	1
City	2	1
Community College	1	1
Child Care Center	1	
Time Req. Parent		
Minimal	7	2
> Twice minimum	7	3

Figure 2. Geographic Distribution of ECEAP First Year Sample Selection Sites



Oregon Business and Education Partnerships Project

Need for data gathering and analysis

The Education and Work program at NWREL conducted a third party evaluation of Oregon's Business and Educational Parterships project (Paul & Owens, 1988). Authorized through the governor's Student Retention Initiative, this program was intended to establish relationships between community businesses and schools with the goal of keeping at-risk students from dropping out of school. Baseline information on dropout rates and the demographic nature of districts with high dropout rates was needed.

Data gathering activities

Data from the existing regional at-risk youth database was used as the baseline information for this project. Poverty indicators for every school district in Oregon were associated with data indicating the numbers and percentages of youth 16-19 falling into the following categories:

- In military service
- Graduated, Employed
- Graduated, Unemployed
- Graduated, Not in the labor force (not seeking employment)
- Not Graduated, Employed
- Not Graduated, Unemployed
- Not Graduated, Not in the labor force (not seeking employment)

Analysis and Sample Results

Since this was a case study evaluation the data were used for descriptive purposes only. Table 8. contains demographic profiles for the 15 Business and Education Partnership sites. The data outlined above were contrasted with the state average for each program site.

Table 8. Demographic Information Regarding Partnership Sites

District Name	Median Family Income	% of Families Below Poverty	For 16-19 Year Olds: % in Schl	Grad Empl	Grad Unempl	Grad NILF	Drop Empl	Drop Unempl	Drop NILF
La Grande	19,106	10	77	11	2	1	3	2	4
Medford	18,677	12	64	12	4	3	7	2	7
Molalla UHS	20,763	9	66	11	3	3	8	3	7
Ontario	16,224	16	74	8	1	2	5	2	7
Tillamook	18,836	10	59	12	5	2	9	6	8
Jefferson SD (Madras)	17,098	15	54	16	0	2	9	4	17
Greater Albany	20,447	14	66	13	4	2	6	3	5
Bend	19,333	8	65	14	3	3	8	4	3
Corvallis	21,019	10	91	3	0	1	2	1	2
Eugene	20,972	11	80	9	1	1	4	2	3
Grants Pass	15,521	17	68	10	2	3	6	4	7
Gresham UHS	24,167	5	70	16	1	2	8	1	3
Klamath Falls	17,100	13	64	10	1	3	5	6	11
Roseburg	19,619	11	65	13	2	3	7	4	6
Salem	19,844	11	63	12	2	4	7	4	10
State Average	18,373	12	67	12	3	4	6	3	5

Local Use of the Regional Database

The Development of District Profiles

In the initial stages of the Database and School Profiling project, NWREL staff developed materials and participated in workshops to assist school staff in developing school *profiles* to guide school improvement efforts. Several states in the Northwest were also initiating efforts at characterizing their districts along a set of common indicators.

Database project staff undertook the task of designing a district level profile with the following features:

- Uniform information on demographic, financial and student achievement characteristics;
- Statewide averages on each indicator for comparison purposes;
- The range and average on each indicator from a cluster of other districts within the state having similar demographic characteristics; and
- Graphic displays of selected indicators.

These profiles could consist of a uniform set of indicators within, but not across, states since the state-specific data obtained differed for each state. This, combined with limitations in project resources suggested the development of the profiling system within a single state for demonstration purposes. Based on requests received, and the quality of its state-specific data, Oregon was selected as the pilot state.

The prototype Oregon district profile is shown in Table 9. Indicators are displayed in the first column in four groups: Comparison Criteria, Local Demographics, School Characteristics and School Outcomes. Moving across the columns, the value of each indicator for the district being profiled is given in the next column, followed by minimum, mean and maximum values for the comparison group of districts, and, finally, statewide averages.

Attached to the profile form were several graphic displays. Each display contained several district indicators, and accompanying statewide averages. The district value was displayed in relation to a "comparison band" constructed around the statewide average (+/- 1 standard deviation in width). An example is shown in Figure 3.

The principle of providing comparative indices for district-level interpretation was a critical one in the development of the profile. Local interest in seeing "how we stack up" to other districts and state averages was a familiar one to project staff.

The method for selecting other districts across the state with "similar characteristics" posed important technical questions. Which indicators should be used to represent these similarities? How closely should districts match before including them in the comparison sample?

The selection of matching indicators had both conceptual and statistical solutions. At a conceptual level, our experience was that, when districts in the Northwest thought of other districts "like themselves", two criteria were first in their minds: size and rurality.

On a more statistical basis, we selected three other indicators which, through factor analysis of fifteen demographic indicators, had two important properties. First, they were relatively independent of each other and, secondly, they accounted for much of the demographic variation among the 1,300 districts throughout the region. The indicators chosen were measures of poverty, adult educational attainment and mobility.

Table 9. Prototype Oregon District Profile

Characteristic	Target District	Similar Districts (n=14) Mean	Min.	Max.	STATE
Comparison Criteria					
Enrollment (FY 85)	775	713	519	976	1,464
% Rural (1980)	100%	100%	100%	100%	80%
% School-aged Children in Poverty	13%	13%	9%	25%	12%
% Families w/ No Workers	14%	14%	9%	16%	13%
% in Same Co. as 1975	76%	73%	59%	80%	69%
Local Demographics (1980)					
% Adults Not HS Grads	31%	31%	22%	38%	28%
% Children w/ Limited English	0.0%	.4%	0.0%	3.6%	0.5%
% One-Parent Families	14%	13%	7%	22%	13%
% Mothers Working	45%	48%	40%	55%	49%
% Minority	2%	3%	0%	8%	3%
Personal Income per Student	$31,272	$29,269	$23,224	$38,794	$34,408
Median Family Income	$17,245	$17,992	$13,912	$21,238	$18,552
% Adults w/ No Children in Household	59%	59%	48%	67%	61%
School Characteristics (FY 85)					
Average Teacher Age	42	41	37	43	40
Average Teacher Experience (years)	12.2	11.3	7.6	13.8	10.9
Average Teacher Salary	$23,964	$21,615	$19,039	$23,964	$21,018
% Teachers w/ Grad. Degree	40%	35%	5%	49%	28%
Student/Teacher Ratio	15.5	17.3	13.0	25.9	15.8
% Transfer in	13%	12%	8%	17%	12%
% Free or Reduced Lunch	42%	38%	18%	57%	36%
Local Budget $ / Student	$3,308	$2,451	$1,279	3,436	$2,698
Total Budget $ / Student	$4,438	$4,042	$3,068	$4,962	$4,369
School Outcome Measures					
% 16-19 Year Olds, Dropouts (1980)	16%	15%	5%	23%	14%
% Not in Labor Force	46%	49%	20%	85%	37%
% in Lowest Quartile (1986):					
Math	11%	16%	10%	23%	13%
Reading	21%	12%	2%	22%	11%

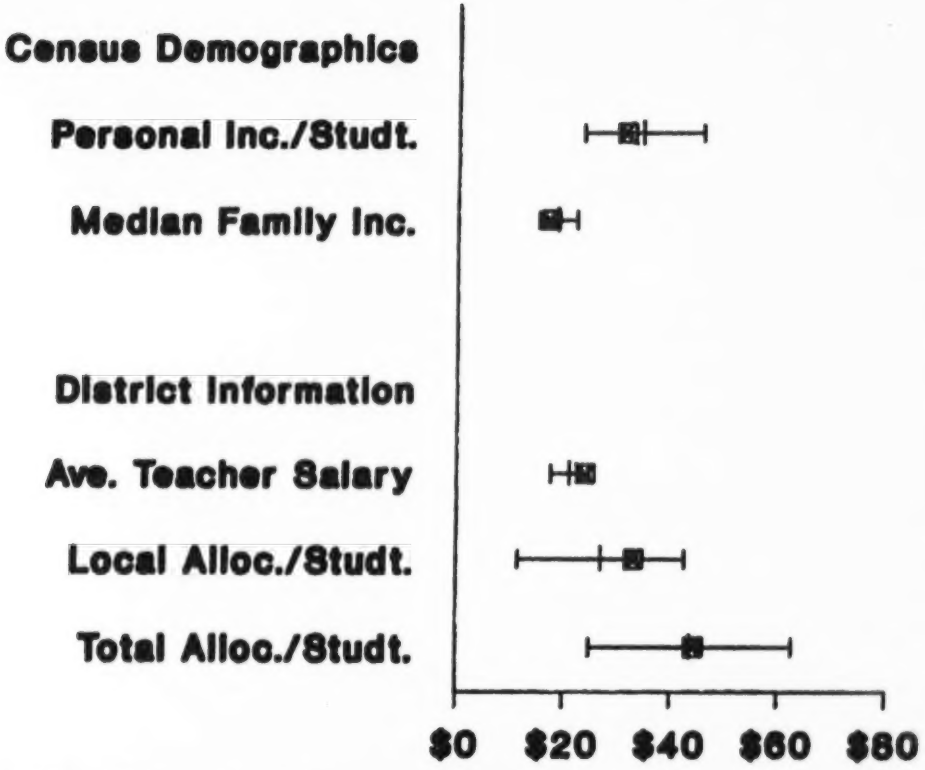
Notes: Similar districts are chosen if they are within a small range of the target district on at least 4 of the 5 matching criteria. Districts must match on both enrollment and rurality.

The state average is an unweighted average of all districts with data.

'M' indicates data is missing or calculated with 0 in the denominator.

Figure 3. School District Profile: Financial Resources

School District Profile
OR S D
Financial Resources



I +/- 1 Std. Dev.

\+ State Mean

□ S D

Allocation figures are in hundreds of dollars. All others are in thousands.

Northwest Regional Educational Laboratory

Together, these five indicators comprised the matching criteria for the profiles. The decision rule as to whether a district was similar enough to the profiled district to be included in the comparison sample was, again, both a conceptual and statistical one. First of all, any district within .5 standard deviations of the district on a given indicator was considered a "match" on that particular indicator. But how many "matched" indicators were required for a district to be considered part of the comparison sample? Ultimately, there were two hurdles:

- A district must match (i.e., be within .5 standard deviation units) the profiled district on both size and rurality; and
- The district must match the profiled district on at least two of the other three comparison indicators

At a conceptual level, to include a district which was not a match on both size and rurality would lose credibility with local staff as they interpreted the profile information. Statistically, our analyses indicated that an insistence on a match on all five indicators would have resulted in a too small a comparison sample of districts in most instances.

Given this careful development, the district profiling system saw many uses in NWREL activities. Laboratory Planning and Service Coordination staff requested profiles for those districts interested in strategic, long range planning activities. They found that an initial look at the status of the district in relation to the comparative indices described above was a useful perception check and stimulant to discussion of long range goals. Staff conducting site visits to selected Oregon districts for evaluation or technical assistance services also profitted from this data-based "snapshot" of district characteristics.

These uses, however, taught project staff valuable lessons in appropriate presentation of statistical data to decision-making audiences with varying degrees of familiarity or sophistication with data collection systems. Some of these were:

Clarify the definitions of the indicators--Some of the indicators on the profile are composites or transformations of several, more fundamental, indicators. The "income per pupil" ratio is a good example. While not a familiar data element to local staff, it made conceptual sense when explained. It was derived to express a sense of the financial ability of the district to support its student population.

Focus on the overall picture the profile provides, not the validity of a single value--Local staff are more familiar with some of the indicators for their districts than were the consultants who prepared the profile. District enrollment is a good example. If the profile indicates the an enrollment figure different from what the local staff know to be the case, the credibility of the entire profile can be damaged. In small districts, these discrepancies can be on the order of five or ten students. Again, by proactively conceding the possible fallibility of any particular data element, while maintaining that, collectively, the indicators likely provide a highly valid picture of the district, these difficulties were averted.

Large samples provide better comparative frames of reference than small ones. This "standard error" principle well known to researchers can make good sense to local decision makers, but is not intuitively obvious. The profile form indicates the number of districts included in the comparison sample. As local staff reviewed the values of all indicators for their district and compared it with those in the comparison sample, anomalies inevitably arose when the sample consisted of only three or four districts. Over time, there will be more variation in these indicators for that group if it is small than if it consists of a relatively large number of districts. Statewide averages provide a more stable, if less contextually specific, comparison.

LESSONS LEARNED AND FUTURE DIRECTIONS FOR THE REGIONAL DATABASE

A demonstration project inevitably includes many false starts, dead ends and inefficient procedures. This one was no exception. In this section of the report, the authors point out some of the primary "lessons learned" in this developmental effort and their implications for the "second generation" regional database.

Difficulties in data gathering

National data sources such as the Census and information available from the Center for Educational Statistics (CES) Common Core Data tapes (CCD) are much more consistent and cost-effective means of gathering such wide scale information than trying to gather information from state and local sources. There are two primary drawbacks to these sources however, the datedness of the information and difficulties with merging data on a district level.

Outdated data. The 1980 Census data has limited usefulness in 1988 and subsequent years for factors that have changed a great deal within the decade. For example, other sources of information indicate that in the Northwest, the numbers of families, and in particular the numbers of children, living in poverty have increased. There have also been dramatic demographic changed in a number of other family related factors. However it will be at least two years before the 1990 census data is available, and the comprehensiveness of the census data is of value.

Merging data from different sources. We have found as much as a 20% mismatch in district identifiers between Census and CCD and current state-wide district listings. This is due to a number of factors such as the following:

- Districts disappear or are consolidated over time;
- District names may be different on different data sources;
- District identifying numbers may be different on different data sources (e.g., the CES number);
- Elementary and secondary school districts may be reported separately on one source and together on another.

For the Rural Education project it was necessary to list names, enrollments and identifying numbers from each source, for every district in each of the seventeen states so that such mismatches could be identified and corrected. The state of Hawaii was excluded from the analyses altogether because, technically, it is a single district and the CCD reports only state-wide data.

We have found it very time consuming and therefore costly to merge data from diverse sources. Information not provided on the national source tapes can be added to the database but this process requires the same type of analysis of mismatches. In addition information for districts and schools is accummulated very differently from state to state. Some states have state-level data, for others it would need to be collected from individual districts. Information gathered from individual states is usually in hardcopy form; very few states we dealt with had achievement data on a computerized system flexible enough to output in the form we needed.

For many applications, such as the Indian Education analysis, it would have been preferrable to have data on a school rather than district level. Consumers of our database efforts have urged us to move in this direction for a number of reasons. Attempting to do this has serious implications for project resources, however, and would greatly magnify the matching issues which were dealt with on the district level.

Possible directions for the future

One solution under consideration for the dated data and costly consistency problems is to use a commerical database (e.g., Quality Education Data, Market Data Retrieval) as the foundation for NWREL regional and special project databases. These lists provide, on both a district and school level, the demographic and financial indicators which have been used in the projects discussed above. Additionally, such information as microcomputer availability and types, updated names of contact persons, ranging from administrators, special education directors and ECIA Chapter project directors to individual teachers are available. Such information could be used for analysis of school resources and staffing or as a selection and mailing list for research surveys

In light of the problems experienced in gathering and merging data such commercial databases look very cost effective because they have already resolved these issues on a school- and district-level basis. In addition the database is updated yearly on such factors as enrollment and school finances. Since the commerical databases are sold to large numbers of consumers the information is ultimately less costly to each consumer. If would not be feasible for NWREL to update information on such a wide scale on a yearly basis. One issue that may not be resolved is the recency of poverty indicators since many commerical databases do rely on census data for some demographic indicators.

The use of commerical data sources can speed the acquisition of data, particulary for new areas such as Early Childhood Education. For example the List Council, Inc. database encompasses 152,948 Early Childhood Education centers including daycare, preschools, Montessori schools, Head Start centers, and public and private pre-Kindergartens and Kindergartens. It specifies such characteristics as bilingual or special education programs, and after school care.

The fundamental source of data for a regional database is only one of several looming issues as the project moves to a "second generation". Others under consideration include the level of specificity of data (school or district), a multi-year component to facilitate the examination of trends over time on selected indicators, the structure of the database (flat, rectangular files suited to statistical analysis vs. a relational format more conducive to on-line searching and queries). These issues are being resolved through a process of needs assessment of an expanding pool of consumers and users of NWREL's regional database.

References

Gabriel, R. & Anderson, P.S. Indentifying At-Risk Youth in the Northwest States: A Regional Database. Northwest Regional Educational Laboratory, March 31, 1987.

Paule, L. & Owens, T. Interim Evaluation Report, Business and Education Partnership Program. Northwest Regional Educational Laboratory, June 30, 1988.

Appendix

Definitions of Indicators Used

in Oregon District Profiles

NWREL DISTRICT PROFILE:
Oregon

The NWREL district profile uses data from the NWREL database. This data is primarily from The 1980 U. S. Census of Population and Housing, with additional information from the Oregon Department of Education (OSDE). The OSDE data is from a variety of sources.

The 1980 Census was conducted in April of 1980. Most information in the district profile refers to the status of people on that date. Income data is based on each person's income for the year of 1979.

Most of the OSDE data is from the School Profile data of FY 1986. This was collected by the state from Fall 1985 school reports and from FY 86 Pupil Participation Accounting Reports. The data was stored by school. We have aggregated this school information to form district figures. Data on free and reduced lunches was collected from a computer printout pertaining to district lunch programs for FY 86 supplied by OSDE.

Enrollment	The OSDE data includes both enrollment in October, obtained from the school reports submitted to the state each fall and Average Daily Membership (ADM), obtained from the Pupil Personnel Accounting Report submitted quarterly. Fall enrollment is chosen for use in the district profile.
% Rural	Percent of the population within the boundaries of the school district living (in 1980) in areas defined as rural by the census bureau. A rural area is an unincorporated area where the population density is less than 1,000 per square mile or a town with less than 2,500 residents.
% School-age Children in Poverty	The percent of children age 5 through 17 (in April 1980) who lived in families where the family income (in 1979) was below the federally defined poverty level.
% Families w/ No Workers	Percent of all families where no person over age 16 was employed in the last week of March, 1980. A family is any household of 2 or more persons related by birth, marriage, or adoption.
% in Same County as 1975	Percent of all persons living in the school district (in 1980) who also lived in the same county five years earlier. This includes people who moved into the district from a neighboring district. A person who moved from Portland to Milwaukee would not be counted as having lived in the same county for No. Clackamas School District. Someone who moved from Oregon City to Milwaukee would.
% Adults Not HS Grads	Percent of all persons 25 years old or older who did not have a high school diploma or equivalent in 1980. One key point is that the people with GED certificates are counted as high school graduates. This figure is based on people living in the school district. They could have gone to school anywhere.

% Children w/ Limited English	Percent of all children 5 through 17 years old who speak only limited English.
% One-Parent Families	Percent of all families with only one adult. Two related adults are counted as a family (in the denominator). Unmarried, unrelated adults sharing a house are not counted as a family. If one of those adults has children at home, they form a one-parent family–even though the household contains two adults.
% Mothers Working	This is simply the percent of women with their own children in their household who were working the last week of March, 1980.
% Minority	Percent of population which is not white. Hispanics are counted separately. Some may be counted as white, most were counted as "other race" and thus are minorities here.
Personal Income per Student	This figure is based on the total of all income reported to the census for 1979, divided by the number of public school students (K-12) living within the school district boundaries. This is a measure of the ability of the community to support the needs of the students living there.
Median Family Income	The Census Bureau ranked the income of all the families in a district. This is the amount of total income reported by the family in the middle. One-half of all families made more, one-half made less. This figure complements the Income per Student figure to give an idea of the economic strength of the voters.
% Adults w/ No Children in Household	Of all households, this is the percent with no children under the age of 18. A high figure would indicate a population dominated with elderly and/or unmarried persons. Both groups are less likely to vote for school levies.
Average Teacher Age	This comes from the SDE school profile. The state collects data on every teacher in each district. Teachers birth years (two digits, eg. 50 for 1950) are aggregated in the state data. To get the average age, the following formula was used: 85-(aggregate birth year / number of teachers).
Average Teacher Experience	Also from the SDE school profile. This is the total experience of all teachers in Oregon plus the total experience in other states, divided by the number of teachers.
Average Teacher Salary	As with teacher experience, this is the aggregate salary divided by the number of teachers in the district.
% Teachers w/ Grad. Degree	The SDE school profile data contains the number of teachers with each of 5 levels of education. This figure is the total number of teachers holding masters or doctorate degrees divided by the number of teachers.

Student/Teacher Ratio	This is the reported enrollment in October 1985 divided by the reported number of teachers. It is unclear how the teachers were counted. Teachers at less than 1.0 FTE may be counted at a full value. If a district has an above-average number of part-time teachers their student/teacher ratio will be artificially deflated.
% Transfer In	In addition to the number enrolled in October, the number of students who transfer into the district during the year are reported in the SDE data. This is data collected from the quarterly Pupil Personnel Accounting Report submitted to the state from each building. The state data used here provides the sum of the four reports on selected items. Transfers are the sum of all transfers into the district. This is divided by the total of all transfers into the district, plus all transfers within the district, plus all students enrolled in the same school for the entire year. Withdrawals are not listed in the data, so the percent who transfer out cannot be calculated.
% Free or Reduced Lunch	The state data (on paper) gave the number of free meals, the number of reduced price meals, and the total number of meals served for each month of FY 86. The percent of free and reduced lunches is the total number of free and reduced lunches served during the year divided by the total of all lunches served.
Local Budget $ per Student	The budget figures came from a Summary of 1984-1985 Audited Resources of Oregon School Districts and ESD's, a report of each district's funds–by 8 sources–plus beginning and ending balances for FY 85. This ratio uses "current year's taxes, "other local and intermediate resources" and "county school fund and federal forest fees" in the numerator. When this ratio was first calculated we used the only current enrollment figures available. The number of students is the Average Daily Membership from FY 86. What source the ADM came from is not clear. Since the ADM figures for FY 86 are very close to the fall enrollment figures for 1985, this ratio has not been re-calculated. Likewise, we have compared ADM figures for FY 85 and FY 86. These, too, are quite similar–within the range of the difference between the fall enrollment and ADM figures for FY 86.
Total Budget $ per Student	This is the same as the above figure, except that all state and federal sources are also included in the numerator. State and federal sources are "basic school support fund", "common school fund", "other state revenue", and "federal revenue".
% 16-19 Year Olds, Dropouts	Percent of all people age 16 through 19 who were not enrolled as of February 1, 1980 in a school and have not completed a high school degree or equivalent.
% Not in Labor Force	Percent of the dropouts included above, who are not and are not looking for work.

% in Lowest Quartile, Math and Reading	In the spring of 1986 the state of Oregon conducted a sample assessment of the achievement of eighth graders. 115 districts participated in the math test, 103 districts gave the reading test. Results of the test were linked to a national norm. The district profile lists the percent of students in the district who scored as low or lower than 25% of the national norm group.

(MOST OF THIS IS AN INTELLIGENT GUESS AS TO THE SOURCE OF THE DATA. FOR INSTANCE, I KNOW THIS IS EIGHTH GRADE DATA AND I KNOW THAT THERE WAS A 33% SAMPLE ASSESSMENT IN THE SPRING OF 1985. I DON'T KNOW OF ANY OTHER STATE ASSESSMENTS, BUT I BELIEVE THAT ALL THE DATA IS FROM FY86, MAYBE IT ISN'T! HOW THE TEST WAS NORMED IS NOT CLEAR, EITHER. THE STATE ASSESSMENT IN 1985 USED A STATE-DEVELOPED TEST, I THINK.)